ACADEMIC PRESS
Available at WWW.MATHEMATICSWEB.ORG POWERED BY SCIENCE DIRECT
Journal of Complexity 19 (2003) 272–285
Journal of COMPLEXITY
http://www.elsevier.com/locate/jco

Some theoretical aspects of generalised quadrature methods

G.A. Evans[a,*] and K.C. Chung[b]

[a] Department of Mathematical Sciences, De Montfort University, Leicester LE1 9BH, UK
[b] Hong Kong Polytechnic University, Hong Kong

Received 6 March 2002; accepted 28 November 2002

Abstract

Generalised quadrature methods rely on generating quadrature rules for given irregular oscillatory weight functions $w(x)$ commonly belonging to the class $C^n[a,b]$, for some usually small n. If these weight functions are known to satisfy $Lw = 0$ for a differential operator L, then Lagrange's identity

$$gLw - wMg = Z'(w,g)$$

can be used to generate a quadrature rule by forcing exactness for a set of basis functions.

Theorems which give conditions under which the computed quadrature rules will yield results correct to a required precision (usually that of the machine being employed) underpin the practical rule, and finite range integrals with weights such as $\sin(q(x))$ and $J_n(q(x))$ have been successfully integrated, for $q(x) \in C^2[a,b]$. Doubly oscillatory weights also become feasible with weights such as $J_n(q_1(x))J_m(q_2(x))$.

More recent work has considered multiple quadratures and the special problems which arise with the commonly occurring infinite range integrations. In the latter case, the direct approach results in violations of the conditions of the underlying theorem and requires some modification for success.

This approach has enabled several diverse practical problems to be attempted including integrals from financial market predictions, from chemical reactor analysis, from coherent optical imaging and from wave analysis on sloping beaches.

*Corresponding author.
E-mail address: gaevans@dmu.ac.uk (G.A. Evans).

0885-064X/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0885-064X(03)00004-9

1. Introduction

Recently, there has been major progress in the development of numerical quadrature methods for the evaluation of highly oscillatory irregular integrands. The regular problem is typified by integrals of the form

$$\int_a^b f(x) \begin{matrix} \sin \omega x \\ \cos \omega x \end{matrix} dx, \tag{1}$$

where ω is large. A quite different philosophy is required for irregular problems compared with the traditional approach of fitting a polynomial to the non-oscillatory part of the integrand, as in [7,11].

The most common irregular problem has the form

$$\int_a^b f(x) \begin{matrix} \sin q(x) \\ \cos q(x) \end{matrix} dx, \tag{2}$$

where $q(x)$ absorbs the large parameter ω. Such integrals arise in applications such as waves on sloping beaches, [3]; in spin calculations in quantum mechanics, [12]; in the calculation of statistical distributions used in financial analysis, [9]; in coherent optical imaging, [2] and in diffraction integrals, [10].

Integrals of the form

$$\int_a^b f(x) J_n(q(x))\, dx, \tag{3}$$

and similar integrals with other Bessel functions form the next most common class. Doubly oscillatory integrands with products of Bessel functions, and products of Bessel and trigonometric functions make for a further complication.

Generalized quadrature methods depend on constructing a relation which involves on one side an integral of an expression containing a general function $f(x)$, and on the other side the value of this integral in terms of f, or a function $v(x)$, where $\mathscr{L}v(x) = f(x)$ where $\mathscr{L}$ is a differential operator and $f \in L^1$. Such an expression arises in [5] in which a second-order differential equation

$$u'' + P_1 u' + P_0 u = 0 \tag{4}$$

is associated with the non-homogeneous equation

$$v'' + P_1 v' + P_0 v = f. \tag{5}$$

Multiplying (4) by v and (5) by u and subtracting, and then setting $z = uv' - u'v$ gives

$$z' + P_1 z = uf. \tag{6}$$

Eq. (6) integrates using the integration factor

$$e^{\int P_1\, dx} \tag{7}$$

to give

$$(uv' - u'v)e^{\int P_1\, dx} = \int u(x) f(x) e^{\int P_1\, dx}\, dx. \tag{8}$$

Hence, we have an exact integral in terms of f in which $u(x)$ can be considered as a weight function which satisfies (4). It is easy, in principle, to generate a quadrature rule based on (8) by using

$$(uv' - u'v)e^{\int P_1\,dx} = \int u(x)f(x)e^{\int P_1\,dx}\,dx = \sum_{k=0}^{N} a_i f(x_i) + E_N^{(1)}(u,f) \tag{9}$$

which can be made exact for a choice of $N+1$ trial functions v_i, possibly $v_i(x) = x^i$. Hence the first and last terms in (9) will generate $N+1$ simultaneous linear equations for the weights a_i, with the assumption that $\int P_1\,dx$ is known analytically. The abscissae x_i also need to be fixed and more will be said on this choice later. Evans and Webster demonstrate the application of this approach to a range of irregular oscillatory integrals, and show the practical capability of the method to high order.

A second such exact quadrature formula is Lagrange's identity

$$gLw - wMg = Z'(w,g), \tag{10}$$

and this is discussed in [4]. By choosing the weight function to satisfy $Lw = 0$, as in (4), and integrating both sides of (10) gives

$$\int_a^b wMg\,dx = -Z(w,g_i)\big|_a^b = \sum_{k=0}^{N} a_k f(x_k) + E_N^{(2)}(w,f), \tag{11}$$

with $Mg = f$, and again (11) is made exact for a choice of trial functions $g = g_i(x)$ to yield self-generated quadrature rules for a given weight function $w(x)$. Two common test sets are used, $g_i(x) = x^i$ and $g_i(x) = T_i(x)$ for $i = 0, 1, \ldots, N$. The weights a_k will therefore satisfy the set of linear algebraic equations

$$\begin{bmatrix} Mg_0(x_0) & Mg_0(x_1) & \cdots & \cdots & Mg_0(x_N) \\ Mg_1(x_0) & Mg_1(x_1) & \cdots & \cdots & Mg_1(x_N) \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ Mg_N(x_0) & Mg_N(x_1) & \cdots & \cdots & Mg_N(x_N) \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} -Z(w,g_0)\big|_a^b \\ -Z(w,g_1)\big|_a^b \\ \vdots \\ \vdots \\ -Z(w,g_N)\big|_a^b \end{bmatrix}. \tag{12}$$

These equations can be written succinctly as

$$\mathbf{Ga} = \mathbf{m} \tag{13}$$

with residue $\mathbf{r}$ defined by

$$\mathbf{r} = \mathbf{Ga} - \mathbf{m}. \tag{14}$$

The approximation to $f(x)$ involved in the quadrature is non-standard, namely

$$f(x) = \sum_{i=0}^{N} a_i f_i(x) + E_N^{(3)}(f, g), \tag{15}$$

where $f_i(x),\ i = 0, \ldots, N$ is the set of functions for which the quadrature is exact. Hence

$$f_i(x) = Mg_i(x), \tag{16}$$

where $g_i(x)$ are the trial functions as before. Several questions arise. First, what is the form of the error term $E_N^{(j)}(u,f)$ in these formulae. Second, what are the criteria for the choice of the abscissae x_i, and third for what range of N will the resulting quadrature formulae be stable. We note at this stage that the question of when the quadrature formulae generated by this process are stable may be a quite different issue from the question of when the underlying linear equations are stable.

Chung et al. [4] demonstrate that the 'computed' weights will satisfy the defining linear equations to yield a residue of the order of the required machine accuracy. As long as this occurs, the corresponding quadrature rule will yield the correct values. This is demonstrated in [6] for the classical Gauss-Legendre quadrature rule. In the present paper some limitations to this process are shown to occur for sufficiently ill-conditioned problems.

2. Stability problems

There are at least two rather separate issues here. One problem is the stability of the process for generating a generalised quadrature rule, which is looked at in this paragraph. The obvious second problem is the stability of the resulting quadrature rule which requires that the generated weights are all positive. This latter question is interlinked with the choice of the abscissae as it is clear from the numerical experimentation that equally spaced points gives rise to weights of varying sign, whereas cosine or Gaussian points give positive weights.

Considering the first problem, practical tests demonstrate that limitations arise when considerations extend beyond the oscillatory integrands of the earlier studies. For example with weight functions with singularities there are quite low values of N beyond which the quadrature rules breakdown. Equally, consideration of integrands of the infinite and semi-infinite ranges have uncovered rules which are generated by linear equations which are so ill-conditioned, that the residue, $\mathbf{r}$, is no longer zero and the quadrature formulae fail.

A quadrature rule is defined by the set of functions for which it is exact. These conditions lead to sets of possibly non-linear equations for the weights and/or the abscissae. In the rules being considered here the abscissae are pre-specified and the algebraic equations are linear. However the weights being generated will be subject to errors due to the ill-conditioning of the defining equations. Hence, the weights being found will be defined as *computed weights* to distinguish them from the *exact*

weights. In other words, the computed weights may have no correct digits compared with the exact weights to the same precision. A practical demonstration of this effect was made in the short paper Evans and Webster [6] in which the classical Gauss–Legendre abscissae were assumed, and linear algebraic equations were solved to get the Gauss–Legendre *computed weights*. As the order increased these computed weights differed more and more from the exact weights, but when used in a quadrature rule gave results to the specified accuracy which were indistinguishable from the results using exact weights. The same phenomenon occurs in the quadrature rules generated in the present paper, and conditions under which the phenomenon occurs are found in the following theorems.

3. Implementation equivalence

The following section answers some of the theoretical aspects of the generalised quadrature methods now in use for oscillatory integrands. These methods use Lagrange's identity given by Eqs. (10) and (11).

There are two methods for the implementation of (11) to evaluate integrals of the form

$$Q(f) \equiv \int_a^b w(x)f(x)\,dx, \tag{17}$$

where $w(x)$ is a given weight function, and $f(x)$ is the argument function.

For each $n = 0, 1, 2, \ldots$, let $x_0, x_1, \ldots, x_N$ (which may depend on N) be the $N+1$ nodes in the interval of integration $[a, b]$. Let $g_0, g_1, g_2, \ldots$ be a sequence of basis functions.

The nodes and the basis functions are given and fixed.

Define the column vectors (each with $N+1$ components):

$$\mathbf{f} = [f_0, f_1, \ldots, f_N]^T \tag{18}$$

and

$$\mathbf{m} = [m_0, m_1, \ldots, m_N]^T, \tag{19}$$

where

$$f_k = f(x_k), \quad m_k = Q(g_k), \quad k = 0, 1, \ldots, N. \tag{20}$$

The first implementation of the generalized rules is an $(N+1)$-point formula of the form

$$Q(f) \simeq \sum_{k=0}^{N} a_k f(x_k) = \mathbf{f}^T\mathbf{a}, \tag{21}$$

where a_k are independent of $f(x)$ and

$$\mathbf{a} = [a_0, a_1, \ldots, a_N]^T. \tag{22}$$

Rule (21) is required to be exact when $f(x)$ is put equal to the basis functions $g_s(x)$ $(s = 0, 1, 2, \ldots, N)$. That is to say, the coefficients a_k are required to satisfy the system of linear equations

$$\sum_{k=0}^{N} a_k g_s(x_k) = m_s, \quad s = 0, 1, 2, \ldots, N \tag{23}$$

which, in matrix form, is

$$\mathbf{Ga} = \mathbf{m}, \tag{24}$$

where $\mathbf{G}$ is the square matrix $[G_{s,k}]$ of order $N + 1$ with $G_{s,k} = g_s(x_k)$. In this method, the moments m_s are known by computing $-Z(w, g_i)|_a^b$, and the coefficients a_k are found by solving the linear system (24) numerically.

The second implementation is to compute $Q(f)$ as follows. Assume that $f(x)$ can be approximated by a finite series expansion of the form

$$f(x) \simeq \sum_{s=0}^{N} b_s g_s(x), \tag{25}$$

where b_s are constants, so that the finite series interpolates $f(x)$ at the nodes x_k $(k = 0, 1, \ldots, N)$. It follows that the coefficients b_s satisfy the linear system

$$\sum_{s=0}^{N} b_s g_s(x_k) = f(x_k), \quad k = 0, 1, 2, \ldots, N, \tag{26}$$

which, in matrix form, is

$$\mathbf{b}^T\mathbf{G} = \mathbf{f}^T, \tag{27}$$

where

$$\mathbf{b} = [b_0, b_1, \ldots, b_N]^T. \tag{28}$$

The linear system (27) is solved for $\mathbf{b}$ and the approximate value of the required integral $Q(f)$ follows using the formula

$$Q(f) \simeq \sum_{s=0}^{N} b_s m_s = \mathbf{b}^T\mathbf{m}. \tag{29}$$

Theorem 1. *The two implementations described above give the same approximate values for the integral $Q(f)$.*

Proof. Using (21) and (29)

$$\mathbf{f}^T\mathbf{a} = (\mathbf{b}^T\mathbf{G})\mathbf{a} = \mathbf{b}^T(\mathbf{Ga}) = \mathbf{b}^T\mathbf{m}. \tag{30}$$

This theoretically common value is denoted by $\hat{Q}(f)$ in the following discussion.

4. Rounding error analysis

In either of the above implementations, a system of $N+1$ linear equations is solved. If N is large, some inaccuracy due to roundoff is to be expected, partly caused by the instability of the coefficient matrix $\mathbf{G}$. Here, the error in the computed integral, arising from the inaccuracy of the solution of the linear system is studied.

In the following theorems, the norms being used are the 2-norms so that Schwarz's inequality can be applied.

Theorem 2. *For the exactness method, suppose that in solving* (24), *the solution is* $\hat{\mathbf{a}}$ *and the relative residual satisfies*

$$r_1 \equiv \frac{||\mathbf{G}\hat{\mathbf{a}} - \mathbf{m}||}{||\mathbf{m}||} < \varepsilon. \tag{31}$$

Then, for absolute error, we have

$$|\hat{Q}(f) - \mathbf{f}^T\hat{\mathbf{a}}| \leqslant ||\mathbf{b}||||\mathbf{m}||\varepsilon, \tag{32}$$

and for relative error, we have

$$\frac{|\hat{Q}(f) - \mathbf{f}^T\hat{\mathbf{a}}|}{|\hat{Q}(f)|} \leqslant \lambda_1\varepsilon \tag{33}$$

where

$$\lambda_1 = \frac{||\mathbf{b}||||\mathbf{m}||}{|\mathbf{b}^T\mathbf{m}|}. \tag{34}$$

Proof.

$$\begin{aligned}
\text{The absolute error} &= |\hat{Q}(f) - \mathbf{f}^T\hat{\mathbf{a}}| \\
&= |\mathbf{b}^T\mathbf{m} - \mathbf{b}^T\mathbf{G}\hat{\mathbf{a}}| \quad \text{(by Theorem 1 and Eq. (24))} \\
&\leqslant ||\mathbf{b}||||\mathbf{m} - \mathbf{G}\hat{\mathbf{a}}|| \quad \text{(Schwarz's inequality)} \\
&\leqslant ||\mathbf{b}||||\mathbf{m}||\varepsilon. \qquad (35)
\end{aligned}$$

The inequality for the relative error follows at once using Theorem 1. □

We have a similar theorem for the expansion–interpolation method.

Theorem 3. *For the expansion–interpolation method, suppose that in solving* (27), *the solution is* $\hat{\mathbf{b}}$ *and the relative residual satisfies*

$$r_2 \equiv \frac{||\hat{\mathbf{b}}^T\mathbf{G} - \mathbf{f}^T||}{||\mathbf{f}||} < \varepsilon. \tag{36}$$

Then, for absolute error, we have

$$|\hat{Q}(f) - \hat{\mathbf{b}}^T\mathbf{m}| \leqslant ||\mathbf{f}||||\mathbf{a}||\varepsilon, \tag{37}$$

and for relative error, we have

$$\frac{|\hat{Q}(f) - \hat{\mathbf{b}}^T\mathbf{m}|}{|\hat{Q}(f)|} \leqslant \lambda_2 \varepsilon \tag{38}$$

where

$$\lambda_2 = \frac{||\mathbf{f}||||\mathbf{a}||}{|\mathbf{f}^T\mathbf{a}|}. \tag{39}$$

Proof.

$$\begin{aligned}\text{The absolute error} &= |\hat{Q}(f) - \hat{\mathbf{b}}^T\mathbf{m}| \\ &= |\mathbf{f}^T\mathbf{a} - \hat{\mathbf{b}}^T\mathbf{G}\mathbf{a}| \quad \text{(by Theorem 1 and Eq. (21))} \\ &\leqslant ||\mathbf{f}^T - \hat{\mathbf{b}}^T\mathbf{G}||||\mathbf{a}|| \quad \text{(Schwarz's inequality)} \\ &\leqslant ||\mathbf{f}||||\mathbf{a}||\varepsilon.\end{aligned}$$

The inequality for the relative error follows at once using Theorem 1. □

The theorems give the possible upper bounds for the relative errors, induced by the relative residuals, of the computed integrals using the two approaches. These bounds are just multiples of the bounds for the relative residuals r_1 and r_2. The multipliers λ_1 and λ_2 are independent of the condition number of the coefficient matrix $\mathbf{G}$, and in practice have been found to be of order unity.

5. Relation to the condition factor

Hence, define a formula to be *exact to relative accuracy* ϵ if

$$\left|\frac{\int_a^b w(x)f(x)\,dx - \sum_{k=0}^{N} a_k f(x_k)}{\hat{Q}}\right| \leqslant \epsilon, \tag{40}$$

where $\epsilon = O(10^{-m})$ so defining the number of digits, m, being worked to. The normalising quantity $\hat{Q}$ is an estimate of the integral value. The following theorem gives conditions when the relative residual is less than machine accuracy, and hence links the accuracy of the resulting quadrature rule to the condition factor for the underlying matrix G by means of the relative residuals in the above two theorems.

Theorem 4. *Let the defining linear equations for a quadrature rule of the above type be* $\mathbf{G}\mathbf{a} = \mathbf{m}$. *Then if the condition factor* $\kappa(\mathbf{G}) < 1/\epsilon$, *for a predetermined relative accuracy* ϵ, *and if the condition*

$$\frac{||(\mathbf{I} + \mathbf{G}^{-1}\delta\mathbf{G})^{-1}||}{||\mathbf{G}||} \sim 1$$

is satisfied, then the residual error for the linear system will be less than the order of the predetermined relative accuracy.

Proof. Let the rounding errors introduced into $\mathbf{G}$ and $\mathbf{m}$ in the solution process be $\delta\mathbf{G}$ and $\delta\mathbf{m}$, respectively. A quadrature formula will yield results with a relative precision ϵ if the conditions of Theorems 2 and 3 hold for trial functions $g_1(x), \ldots, g_N(x)$. We require, therefore, the relative residual vector $\mathbf{r}$ to be less than the relative accuracy ϵ.

Hence for the computed weights, the result is $\mathbf{a} + \delta\mathbf{a}$ and the residue $\mathbf{r}$ satisfies

$$\mathbf{r} = \mathbf{G}(\mathbf{a} + \delta\mathbf{a}) - \mathbf{m}. \tag{41}$$

The perturbed problem satisfies

$$(\mathbf{G} + \delta\mathbf{G})(\mathbf{a} + \delta\mathbf{a}) = (\mathbf{m} + \delta\mathbf{m}) \tag{42}$$

or

$$\delta\mathbf{G}\mathbf{a} + \mathbf{G}\delta\mathbf{a} + \delta\mathbf{G}\delta\mathbf{a} = \delta\mathbf{m}. \tag{43}$$

The residue $\mathbf{r}$ then becomes

$$\begin{aligned}\mathbf{r} &= \mathbf{G}(\mathbf{a} + \delta\mathbf{a}) - \mathbf{m}\\ &= \mathbf{G}\delta\mathbf{a} = \delta\mathbf{m} - \delta\mathbf{G}\mathbf{a} - \delta\mathbf{G}\delta\mathbf{a}.\end{aligned} \tag{44}$$

Taking second norms then gives

$$\begin{aligned}\frac{||\mathbf{r}||}{||\mathbf{m}||} &\leqslant \frac{||\delta\mathbf{m}||}{||\mathbf{m}||} + \frac{||\delta\mathbf{G}\mathbf{a}||}{||\mathbf{m}||} + \frac{||\delta\mathbf{G}\delta\mathbf{a}||}{||\mathbf{m}||}\\ &\leqslant \frac{||\delta\mathbf{m}||}{||\mathbf{m}||} + ||\mathbf{a}||\frac{||\delta\mathbf{G}||}{||\mathbf{m}||} + \frac{||\delta\mathbf{G}||||\delta\mathbf{a}||}{||\mathbf{m}||}\\ &\leqslant \epsilon_1 + ||\mathbf{a}||\epsilon_2 + K.\end{aligned} \tag{45}$$

The terms ϵ_1 and ϵ_2 are assumed to be of the order of ϵ, as these terms are governed by the representational errors $\delta\mathbf{m}$ and $\delta\mathbf{G}$, and $||\mathbf{a}||$ is bounded. Hence to get our result we need K to also be less than ϵ. But

$$\delta\mathbf{a} = (\mathbf{G} + \delta\mathbf{G})^{-1}(\delta\mathbf{m} - \delta\mathbf{G}\mathbf{a}) \tag{46}$$

as long as $(\mathbf{G} + \delta\mathbf{G})$ is non-singular. Then the third term K in (45) becomes

$$\begin{aligned}K &= \frac{||\delta\mathbf{G}||||\delta\mathbf{a}||}{||\mathbf{m}||}\\ &\leqslant \frac{||\delta\mathbf{G}||||(\mathbf{G} + \delta\mathbf{G})^{-1}||(||\delta\mathbf{m}|| + ||\delta\mathbf{G}||||\mathbf{a}||)}{||\mathbf{m}||}\end{aligned} \tag{47}$$

$$\leqslant \frac{||\delta\mathbf{G}||||\mathbf{G}^{-1}||||(\mathbf{I}+\mathbf{G}^{-1}\delta\mathbf{G})^{-1}||(||\delta\mathbf{m}||+||\delta\mathbf{G}||||\mathbf{a}||)}{||\mathbf{m}||}$$

$$= \kappa(\mathbf{G})||\delta\mathbf{G}||\frac{||(\mathbf{I}+\mathbf{G}^{-1}\delta\mathbf{G})^{-1}||}{||\mathbf{G}||}\left[\frac{||\delta\mathbf{m}||}{||\mathbf{m}||}+\frac{||\delta\mathbf{G}||||\mathbf{a}||}{||\mathbf{m}||}\right]$$

$$\leqslant \kappa(\mathbf{G})\epsilon^2\frac{||(\mathbf{I}+\mathbf{G}^{-1}\delta\mathbf{G})^{-1}||}{||\mathbf{G}||}, \tag{48}$$

where $\kappa(\mathbf{G})$ is the conventional condition factor, $||\mathbf{G}||||\mathbf{G}^{-1}||$.

The classical assumption in the analysis of condition factors, such as in [8], is that $||\mathbf{G}^{-1}\delta\mathbf{G}||$ is negligible, so that the Neumann series for $(\mathbf{I}+\mathbf{G}^{-1}\delta\mathbf{G})^{-1}$ leaves the condition factor κ as the multiplying factor. Hence the condition that

$$\frac{||(\mathbf{I}+\mathbf{G}^{-1}\delta\mathbf{G})^{-1}||}{||\mathbf{G}||}\sim 1 \tag{49}$$

is necessary for the theorem. Eqs. (45) and (48) together with condition (49) yield the required condition that $\kappa(\mathbf{G})<\frac{1}{\epsilon}$.

All the quantities involved are computable, as $||\mathbf{G}|| = \sigma_1$ and $||\mathbf{G}^{-1}|| = 1/\sigma_N$, the largest and smallest singular values of $\mathbf{G}$, respectively. For all the cases which arise in the quadrature context, the value of $\kappa(\mathbf{G})$ is large and this is due to either σ_1 being large or σ_N being small, or both. In the test examples, σ_N has always been order 1 or less, and so justifies the Neumann series assumption. Also σ_1 has never been small. Clearly if σ_1 is large even higher $\kappa(\mathbf{G})$ may be tolerated, so generating enhanced stability which is seen in some of the ensuing examples.

Hence Theorems 2 and 3, together with Theorem 4 give conditions under which the computed weights will yield quadrature results to a predetermined relative accuracy (normally the working machine precision).

6. Practical demonstrations

These effects are well illustrated in practice by considering typical examples from the weight functions which have found common application. The first example typifies the trigonometric cases and is taken from Evans and Webster [6], example I_3, namely

$$\int_0^1 \sin x\cos(500(x^2+x))\,dx. \tag{50}$$

The accurate value for this integral computed by other means is 4.4985939784014(−4). The values of κ, σ_1 and σ_n are tabulated in Table 1. The results were also confirmed using quadruple precision computations, which reproduced to 14 figures the final column.

Table 1

N	σ_1	σ_N	κ	Computed integral approximation
4	1.9(3)	4.6(2)	4.1	4.60030(−4)
8	2.9(3)	2.9(1)	1.0(2)	4.598611114(−4)
16	4.6(3)	4.9(−3)	9.6(5)	4.5985939789914(−4)
32	7.5(3)	9.2(−4)	8.2(6)	4.5985939784014(−4)
64	1.2(4)	6.4(−6)	1.9(9)	4.5985939784014(−4)

Table 2

N	σ_1	σ_N	κ	Computed integral approximation
4	1.7(4)	1.4(3)	1.2(1)	5.35305(−3)
8	5.5(4)	8.6(1)	6.4(2)	−2.076686(−2)
16	1.1(6)	7.5(−1)	1.4(6)	−2.0389271311633(−2)
32	2.2(7)	2.8(−2)	7.9(8)	−2.0388954040983(−2)
64	4.8(8)	2.2(3)	2.2(5)	−2.0388954040985(−2)

Trigonometric examples of this type and for a wide range of $q(x)$ exhibit κ values less than $1/\varepsilon$ and have a value of σ_1 which is large, further helping the stability position. Hence the computed weights satisfy Theorem 4 and we observe accurate integration values.

As a second example, a Bessel function weight was employed in the integral

$$\int_{0.5}^{1} e^x J_{100}(130\cos x)\,dx \tag{51}$$

which is one of the examples in [4], and these results are shown in Table 2.

Now we see that the value of σ_1 is rising and providing increased accuracy over and above the $1/\varepsilon$ effect. Again the accurate value is approached successfully, and the use of quadruple precision yields the same final column to 14 figures except for the final $-2.0388954040985(-2)$ for $N = 64$ which becomes $-2.0388954040983(-2)$.

For the third case, an integral over $[0, \infty]$ is considered with a weight function $e^{-\beta x}$. This integral arose in current studies involving the extension of the trigonometric and Bessel work to the semi-infinite range. The integral was attempted in [1] where the integral was split into two. Here it is sufficient to see the results for one of these integrals, namely

$$\int_{1}^{\infty} \frac{e^{-1/x}\sin(10x)}{x^{1/2}}\,dx \tag{52}$$

which are shown in Table 3.

This is a less stable case as σ_N becomes very small and the enhanced effect no longer applies. Hence the accurate value is approached around $N = 32$ and then a less accurate value occurs at $N = 64$ as the loss of figures arises as Theorem 4 is

Table 3

N	σ_1	σ_N	κ	Computed integral approximation
4	5.1(1)	1.7(−1)	3.0(2)	−1.19094446(−2)
8	8.1(1)	7.7(−5)	1.1(6)	−1.1884629657(−2)
16	1.2(2)	2.5(−6)	5.0(7)	−1.1884680093413(−2)
32	1.9(2)	6.7(−8)	2.9(9)	−1.1884680110436(−2)
64	5.5(2)	6.3(−7)	8.7(8)	−1.1884680094608(−2)

violated. This is confirmed in quadruple precision with the computed value for $N = 64$ being $-1.1884680110448(-2)$.

For this problem, the weight function used is $w(x) = \exp(iq(x) - \beta x)$ and hence the operator Lw is

$$Lw = \frac{dw}{dx} - (iq'(x) - \beta)w = 0 \tag{53}$$

with the adjoint

$$Mz = -\frac{dz}{dx} - (iq'(x) - \beta)z = f \tag{54}$$

and the non-linear concomitant Z is $Z = wz$.

The object of the constant β is that for integrals with no exponent part, a compensating $e^{\beta x}$ can be included in the function $f(x)$ and β can be set very small. Hence the convergence of the monomials is maintained without strongly influencing the function $f(x)$.

A fourth example is based on the logarithmic weight and uses the integral

$$\int_2^\infty \frac{\ln x}{(1+x^2)}\,dx, \tag{55}$$

and the results here are shown in Table 4, where both double and quadruple results are shown.

In this fourth example the rise in κ is entirely due to σ_N getting very small and there is now a marked loss of significant figures with just 7 correct figures from a working accuracy of around 14 figures.

This example is one where the underlying operator is inhomogeneous. The weight is $w(x) = e^{-\beta x}\ln x$ with the operator Lw given by

$$Lw = \frac{dw}{dx} + \beta w(x) = \frac{e^{\beta x}}{x} = h(x). \tag{56}$$

Hence $Mz = -z' + \beta z$ and $Z = zw$. Hence the integral of (10) will now not only involve Z on the right-hand side but also an integral of $h(x)g_i(x)$ from a to b, where $g_i(x)$ is the trial function set. In this test the trial functions were also non-standard being inverse powers of x, so that the equivalent of (11) was

$$\int_a^b wMw_i\,dx = -Z(w, g_i)|_a^b + \int_a^b h(x)g_i(x)dx. \tag{57}$$

Table 4

N	σ_1	σ_N	κ
4	4.2(−1)	5.6(−7)	7.5(5)
8	5.3(−1)	3.5(−8)	1.5(7)
16	6.8(−1)	3.6(−9)	1.8(8)
32	9.1(−1)	6.0(−9)	1.5(8)
64	1.2	4.1(−11)	3.0(+10)

N	Computed integral double precision	Computed integral quadruple precison
4	−6.0970468(2)	−6.097046759(2)
8	6.81255490(−1)	6.8125548914(−1)
16	8.08590816101(−1)	8.08598392456(−1)
32	8.08590919478(−1)	8.085983912467 91(−1)
64	8.08590902364(−1)	8.085983912467 91(−1)

For $g_i(x) = 1/x^{i-1}$ the integral is an exponential integral $E_{i+1}(\alpha\beta)/a^i$ which can be evaluated using a standard approximation.

Care must be taken in interpreting the accuracy of the integrals themselves as three factors are at play. The usual truncation error will be the leading error for small N. As N increases, the truncation error will fall off and the two stability errors will arise. One comes from the inaccuracy in computing the weights and is the subject of this paper. The other arises in the use of the quadrature rule and does not arise if all the weights are positive. This is not necessarily the case in these examples, and the size of any terms with alternating signs in the resulting weights can be monitored. This effect is marked in the fourth example.

The choice of abscissae in the generalised formulae will affect this latter stability. Formulae with equally spaced abscissae will exhibit Newton–Cotes like large alternating weights as, for small ω, the generalised rules will reduce to equivalent Newton–Cotes rules. The cosine weights used in the finite range integrals (and transformed for use on the infinite range), will tend to the Clenshaw rules as ω gets small but there is no guarantee of positive weights in general.

7. Conclusions

The validity of the generalised quadrature rules has been achieved, and results in some practical considerations which can be built into the implementation. It would seem that the following additional codes should be included with a view to confirming the conditions for the resultant quadratures.

(i) Compute the condition factor $\kappa(\mathbf{A})$. (ii) Compute the number of sign changes in the resulting computed weights. (iii) Determine the weights with the smallest and largest modulus.

Condition (i) will confirm the expected loss of figures based on Theorem 4 for the computed quadrature rule. Conditions (ii) and (iii) will give an indication of the stability of the resulting rule to cover the absence of a criteria for the choice of the abscissae. If equally spaced abscissae were used, this would result in large Newton–Cotes like alternating coefficients and quadrature instability, which would be indicated by the violation of the latter two conditions.

References

[1] M. Blakemore, G.A. Evans, J. Hyslop, Comparison of some methods for evaluating infinite range integrals, J. Comput. Phys. 22 (1976) 352–376.
[2] M. Born, E. Wolfe, Principles of Optics, Pergamon Press, New York, 1980.
[3] U.T. Ehrenmark, Far-field asymptotics of the two dimensional linearised sloping beach problem, SIAM J. Appl. Math. 47 (1987).
[4] G.A. Evans, K.C. Chung, J.R. Webster, A method to generate generalized quadrature rules for oscillatory integrals, Appl. Numer. Math. 34 (2000) 85–93.
[5] G.A. Evans, J.R. Webster, A high order progressive method for the evaluation of irregular oscillatory integrals, Appl. Numer. Math. 23 (1997) 205–218.
[6] G.A. Evans, J.R. Webster, The accuracy of solutions of linear equations in practice, Int. J. Math. Educ. Sci. Technol. 29 (1998) 105–112.
[7] L.N.G. Filon, On a quadrature formula for trigonometric integrals, Proc. Roy. Soc. Edinburgh 49 (1928) 38–47.
[8] L. Fox, Introduction to Numerical Linear Algebra, Oxford University Press, Oxford, 1968.
[9] G. Fusai, A double integral for the price of Asian average options, Universita di Novara, 2000.
[10] D.R. Hartree, The evaluation of a diffraction integral, Proc. Cambridge Philos. Soc. 50 (1954) 567–574.
[11] T.N.L. Patterson, On high precision methods for the evaluation of Fourier integrals with finite and infinite limits, Numer. Math. 27 (1976) 41–52.
[12] J.H. Sampson, G.A. Evans, Symmetry reduction of Fourier kernels, J. Comput. Phys. 142 (1998) 109–122.